Primary Sjögren's Syndrome

Delli, Konstantina

Chapter 3A

Diagnostic properties of ultrasound of major salivary glands in Sjögren's Syndrome: a meta-analysis

Konstantina Delli[1], Pieter U Dijkstra[1,2], Alja J Stel[3], Hendrika Bootsma[3], Arjan Vissink[1], Frederik KL Spijkervet[1]

Best Original Article Award
for 2015 In Oral Diseases

Affiliations:

1. Department of Oral and Maxillofacial Surgery

2. Department of Rehabilitation, Center for Rehabilitation

3. Department of Rheumatology and Clinical Immunology

University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Edited version of: Oral Dis. 2015 Sep;21(6):792-800.

Abstract

Objective: To perform a systematic review and meta-analysis on studies examining the properties of ultrasonography of major salivary glands for diagnosing Sjögren's Syndrome.

Materials & Methods: We searched for literature on eight databases. The quality of included articles was assessed with the QUADAS-2 tool. Publication bias, pooled sensitivity, specificity, diagnostic odds ratio and 95% confidence intervals (95%CI) were calculated. Meta-regression analysis was performed.

Results: We identified 37 studies and 33 ultrasonographic scoring systems. High risk of bias was observed in 'patient selection', 'conduct and interpretation of ultrasound' and 'flow of patients and timing of tests' in 78%, 70% and 51% of the studies. We included 29 studies in the meta-analysis. Publication bias was highly probable. Pooled sensitivity was 0.69 (95%CI: 0.67-0.71), specificity 0.92 (95%CI: 0.91-0.93) and diagnostic odds ratio 33.89 (95%CI: 20.75-55.35). Significant heterogeneity was detected between studies. Meta-regression analysis showed that studies with high risk of bias in 'conduct and interpretation of ultrasound' and studies evaluating only parenchymal homogeneity had higher log diagnostic odds ratio (1.09 and 2.49 respectively, $p<0.05$).

Conclusions: The quality of current studies is low thus not allowing to judge the likelihood of salivary gland ultrasonography as a reliable and practical tool in diagnosing Sjögren's Syndrome.

Introduction

Sjögren's syndrome (SS) is one of the most common rheumatic diseases, with a prevalence of 0.05% in the total population. It commonly affects the salivary and lacrimal glands, resulting in a sensation of dry eyes (keratoconjunctivitis sicca) and dry mouth (xerostomia) [1]. Although the exact pathogenic mechanism has not been elucidated, in patients with SS the minor and major salivary glands are characteristically infiltrated by mononuclear lymphoid cells [2]. Enlargement of the major salivary glands, especially the parotid and submandibular is also common. Swelling of the salivary glands is usually bilateral, may be non-painful to slightly tender and intermittent to persistent in nature. The development of lymphomas in SS, in most cases in the parotid gland, can lead also to more persistent unilateral glandular enlargement.

In the assessment of salivary gland involvement in SS, ultrasonography of the major salivary glands merits special attention as a non-invasive, inexpensive, widely available, easily accessible and non-irradiating imaging modality [3]. Ultrasonography enables visualization of deep structures of the body by recording the reflections or echoes of ultrasonic pulses directed into the tissues. Frequencies ranging from 1.6 to 22 MHz are used for diagnostic imaging. B-mode is the most widely used ultrasonography mode; the gray scale of the image consists of pixels whose brightness depends on the intensity of the echo that is received from the corresponding location in the body [4].

Recently, a meta-analysis was published regarding the diagnostic properties, sensitivity, specificity and diagnostic odds ratio of ultrasonography in diagnosing SS [5]. However, the data in the tables in that study did not appear to correspond with the data from the source publications. There seems to be a discrepancy between the data shown in the meta-analysis and the data presented by the source studies [6-11]. Hence, the diagnostic properties of ultrasonography in diagnosing SS remain unclear [12].

The primary objective of this study was to conduct a systematic review and meta-analysis on studies examining the diagnostic properties of ultrasonography of major salivary glands in diagnosing SS in comparison to criteria applied in the classification of SS or in comparison to sialography, scintigraphy, sialometry and histology of the salivary glands. The secondary objective of the study was to evaluate effects of research methodology on diagnostic properties of ultrasound in SS found in the studies.

Materials and methods

Study identification

We conducted a literature search of eight electronic databases (six literature databases and two control trials registries). According to the syntax rules of each database, key words and their combinations were used to identify studies published till June 2014. No restrictions were applied (supplementary Table 1).

Study eligibility

Two observers (K.D. and P.U.D.) independently assessed titles and abstracts identified in the initial search. Inclusion criteria were studies examining the diagnostic properties of ultrasonography of major salivary glands in diagnosing SS in comparison to diagnostic criteria or in comparison to sialography, scintigraphy, sialometry or histology of the salivary glands. Exclusion criteria for titles and abstracts were the following: case reports, case series with fewer than 10 cases, experts' opinions, letters to the editor, review articles, studies that did not report the diagnostic properties of ultrasonography and congress abstracts. If the title and abstract provided limited information or in case of doubt, the studies were moved to the next round (full text assessment). The results of the assessment were compared, and any disagreement was resolved through consensus.

Full texts of the included titles and abstracts were independently assessed according to aforementioned criteria by the same observers. Additionally, the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool was used to assess the risk of bias of the included studies [13]. This tool consists of four domains covering 'patient selection', 'index test', 'reference standard', and 'flow of patients through the study and timing of the index test and reference standard'. The observers (K.D. and P.U.D.) independently evaluated the studies. After each stage of selection, inter-observer agreement was calculated as Cohen's kappa and percentage of agreement. Studies written in a language, which the assessors were not proficient, were translated into English by researchers fluent in both that language and English. These studies (4 Italian, 2 Spanish, 1 Russian, 1 Chinese) were assessed for inclusion criteria and risk of bias by one observer (K.D.).

Data extraction was performed (by K.D.) on study and patient characteristics, and on the diagnostic performance of ultrasonography. Information about ultrasonographic criteria for diagnosing Sjögren's syndrome, type of reference standard, ultrasonographic transducer used, and true positive, true negative, false positive, false negative and accuracy results were collected, using a standardized form. If one cell was empty, in the two-by-two tables 0.5 was added to all cells to be able to calculate a diagnostic odds ratio. If a control group was not included in the study, only true positive and false negative results were collected.

The reporting of this study complied with the 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses' (PRISMA) statement [14] and the 'A MeaSurement Tool to Assess systematic Reviews' (AMSTAR) recommendations [15].

Statistical analysis

Inter-observer agreement was calculated with IBM SPSS Statistics 20 (SPSS, Chicago, Illinois, USA). Publication bias was assessed by plotting the log odds ratio against its standard error using the software Comprehensive Meta-Analysis, Version 3 (CMA, Biostat, Englewood, NJ 07631, USA). Pooling of sensitivity, specificity and diagnostic odds ratio was performed with MetaDisc, Version 1.4 (Hospital Universitario Ramon y Cajal, Madrid, Spain). For the pooled diagnostic odds ratio, a random effects model was used. To analyze sources of heterogeneity between studies, a meta-regression analysis (random effects model) was performed with criteria used to diagnose SS, publication year, type of glands examined and ultrasonographic characteristics assessed with ultrasonographic scoring systems, risk of bias in 'patient selection', 'conduct and interpretation of index test', and 'flow of patients and timing of tests' as predictors for differences in diagnostic properties of ultrasound in SS found in the studies (Comprehensive Meta-Analysis, Version 3).

Table 1: Characteristics of the studies included.

AECG: American European Consensus Group Criteria [47]; C: Criteria; CC: Copenhagen criteria [51]; disease: control group with a non-autoimmune salivary gland disease, not related to dry mouth, but potentially causing abnormal ultrasonographic imaging (e.g. salivary gland tumors); ECSG: European Community Study Group Criteria [48]; FC: Fox Criteria [49]; H: Histology; JDC: The Sjögren's Disease research Committee of the Japanese Ministry of Health and Welfare 1977 Diagnostic Criteria for definite Sjögren's Syndrome [52]; ns: not specified; pSS: primary Sjögren's Syndrome; pt: patient; RJDC: Revised Japanese Diagnostic Criteria [50]; SC: scintigraphy; SL: sialography; sSS: secondary Sjögren's Syndrome; TC: Takashima Criteria [43]; y: years

* Carotti et al., 2001 [19] and Salaffi et al., 2000 [40] used the same study population (patient and control group)

Author	Year	Diagnostic Criteria	Reference standard	Ultrsound transducer (MHz)	SS patients (n) included patients (mean age ± SD, range)			Control (n) included subjects (mean age ± SD, range)				Duration of symptoms/ mean ± SD, (range)
					pSS	sSS	ns	sicca	disease	healthy	ns	
Diederich et al.	1987	ns	H,SL				5					
De Clerck et al.	1988	FC	C,SL	7.5			16	26				
Kawamura et al.	1990	JDC	C	7.5	15	9			7	23		
Akin et al.	1991	H	H	5			5					
Corthouts et al.	1991	FC	C	7.5			16	13				
De Vita et al.	1992	CC	C	7.5	27	26		26		64		
Takashima et al.	1992	TC	C	7.5-10	13 (47, 17-63)	17 (47, 17-63)				10 (45, 22-58)		
Ariji et al.	1996	FC	C	7.5			25	19 (55±11, 29-69)		72 (50±18, 17-88)		
Makula	1996	ECSG	C	7			62 (53.2, 29-80) 53 definite 9 probable		44	25		9.8 y
Yoshiura et al.	1997	JDC	C	7.5			24		19	22		
Makula et al.	2000	AECG	C	7			44	11	27	14		
Salaffi et al.	2000	ECSG	C	7.5	30 (54, 29-75)			30 (53, 30-78)				7.6y (0,5-11) 6.9y (8mo-10y)
Carotti et al.*	2001	ECSG	C	7.5	30 (54, 29-75)			30 (53, 30-78)				7.6 y
El Miedany et al.	2004	AECG	C,H	7	47 (54.8,±5.5, 47-66)			20 (55.6±7.3)			20 (55.6±7.3)	
Niemela et al.	2004	AECG	C	9.6-11	27			27		27		
Hocevar et al.	2005	AECG	C	5-12			68				150	
Chikui et al.	2006	RJDC	C	7			91				41	
Decuzzi et al.	2006	ns	C,SC	7.5			20 (52±5, 35-65)					
Shimizu et al.	2006	RJDC	C	6-14			48 (53.5)				32 (53.5)	
Hocevar et al.	2007	AECG	C	5-12			28 (56.6±11.2, 32-84)			29 (56.7±12.3, 32-78)		
Poul et al.	2008	AECG	C	ns	36 (60, 20-85)	9 (60, 20-85)		15 (60, 20-85)				
Salaffi et al.	2008	AECG	C	7.5-10	77 (54±12.1, 30-78)			79 (53±12.3, 24-81)				-pSS: 2.9y (6mo-11y) -sicca non-SS: 2.8y (4mo-12y)
Shimizu et al.	2008	RJDC	C	8			43 (53.1, 17-80)				29 (53.1, 17-80)	
Wernicke et al.	2008	AECG	C	5-10	57	33		78		148		
Chikui et al.	2009	RJDC	SL	7			89 (52.6±16.4)				103 (55.9±16.2)	
Milic et al.	2009	AECG	C	4-10	107 (54.1, 21-78)						28 (54.1, 21-78)	
Milic et al.	2010	AECG	C	4-10	115	44		50		36		
Obinata et al.	2010	RJDC	C	5-12			36 (48, 13-68)	37 (48, 13-68)				
Tagaki et al.	2010	AECG	C	10			188 (56±13)				172 (55±16)	
Milic et al.	2012	AECG	C	4-10	140 (54.5, 21-78)						50 (52.6, 27-70)	
Cornec et al.	2013	AECG	C,H		78 (56.8± 12,7)			80 (56.8± 12.7)				6.7±7 y
Theander et al.	2014	AECG	C	6-18	105 (61±14.9, 20-91)			19 (57±15, 25-91)	20 (57±15, 25-91)			

Results

Study identification and selection

A total of 1245 papers were identified. After excluding duplicates, 742 papers were retrieved and screened by title and abstract (Figure 1). Subsequently, 700 titles and abstracts were excluded (a list of all identified papers and excluded papers not presented in this paper can be requested from the corresponding author). Cohen's Kappa agreement was 0.72 and overall agreement was 93%. We screened the full text of 42 studies. Finally, 37 studies were included for quality assessment (Figure 1) [3,6-9,11,16-46].

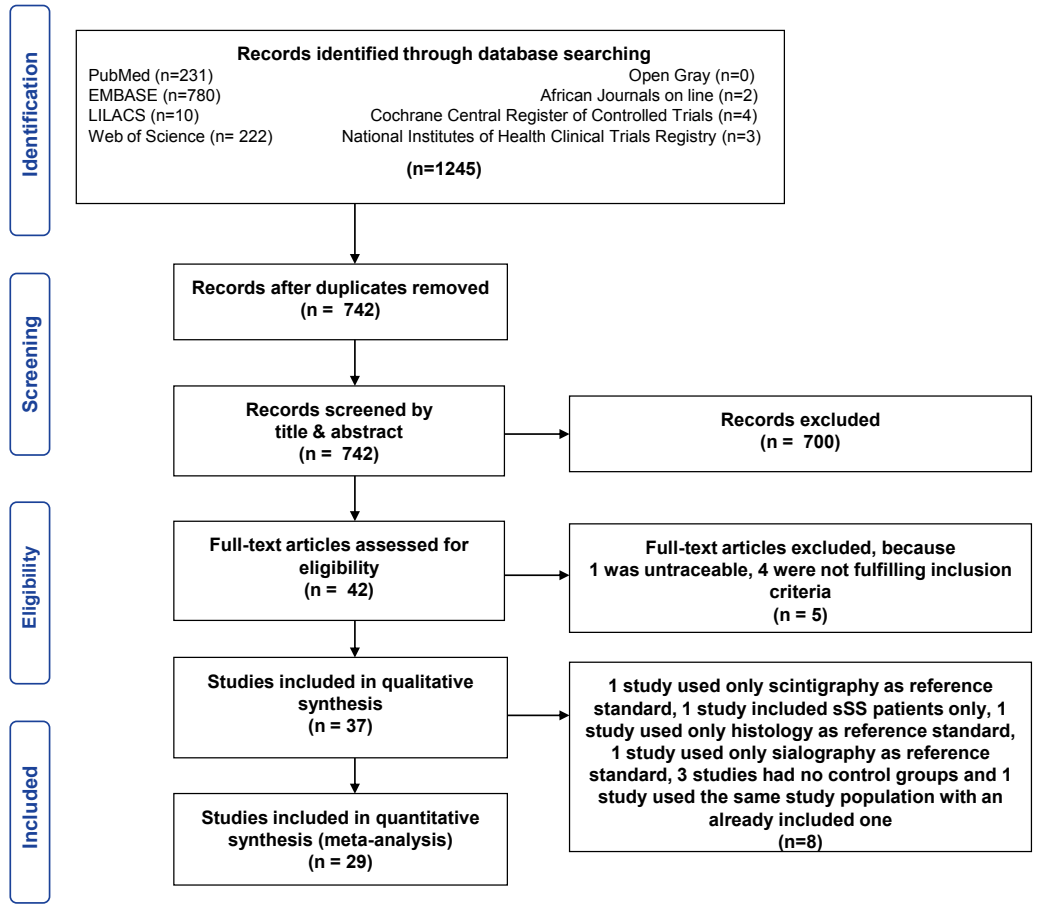
Quality assessment of studies

High risk of bias was observed in 'patient selection', 'conduct and interpretation of index test', and 'flow of patients and timing of tests' (78%, 70% and 51% of the included studies, respectively), while unclear risk of bias was found in the 'conduct and interpretation of the reference test' in 73% of the studies (supplementary Table 2; Figure 2). Kappa and overall agreement at this stage were 0.77 and 86%, respectively. Disagreements were resolved by discussion.

Study characteristics

In the 37 assessed studies, SS was generally diagnosed according to American European Consensus Group Criteria (AECG) [47], European Community Study Group Criteria (ECSG) [48], Fox Criteria (FC) [49], and Revised Japanese Diagnostic Criteria (RJDC) [50]. AECG was used in 15 studies (41%), ECSG in 5 studies (14%), FC in 5 studies (14%) and RJDC in 4 studies (11%). Less common criteria were used in 8 studies (22%) (Table 1) [43,51,52]. The diagnostic properties of ultrasonography were compared to criteria in 30 studies (81%), while in the rest of the studies ultrasonography was compared to sialography, scintigraphy or histology. In total, 2084 patients with SS and 1783 control subjects were included. Nine studies included only primary SS (pSS) patients, 10 studies both pSS and secondary (sSS) patients and 1 study only sSS patients. In the remaining 17 studies the type of the disease (pSS or sSS) was not specified. In 13 studies a healthy control group was included, and in 17 studies a sicca non-SS control group, while in 11 studies no details about the control group were provided. In 6 studies a control group was included with a non-autoimmune salivary gland disease, which was not related to dry mouth, but could cause an abnormal ultrasonography image of the glands (e.g., salivary gland tumors). Overall, 12 studies had multiple control groups and 3 studies no control group.

Figure 1: Study identification and selection progress.



In total 33 ultrasonographic scoring systems were used in 37 studies (supplementary Table 3). This heterogeneity within the scoring systems was related to the type of salivary glands examined, the ultrasonographic characteristics evaluated and the cut-off points applied. Of the scoring systems, 19 were related to both parotid and submandibular salivary glands, 11 were related to only the parotid glands and in 3 systems the type of the examined gland was not clearly mentioned. The main ultrasonographic characteristics assessed were the echogenicity and homogeneity of the glandular parenchyma. Both echogenicity and homogeneity were evaluated in 22 systems, only echogenicity in 8 and only homogeneity in 3 scoring systems. The size of the gland and the clearness of the border were also assessed in 12 and 11 scoring systems, respectively.

Figure 2: Percentage of studies included in the qualitative analysis with low, high, or unclear concern regarding applicability and risk of bias.

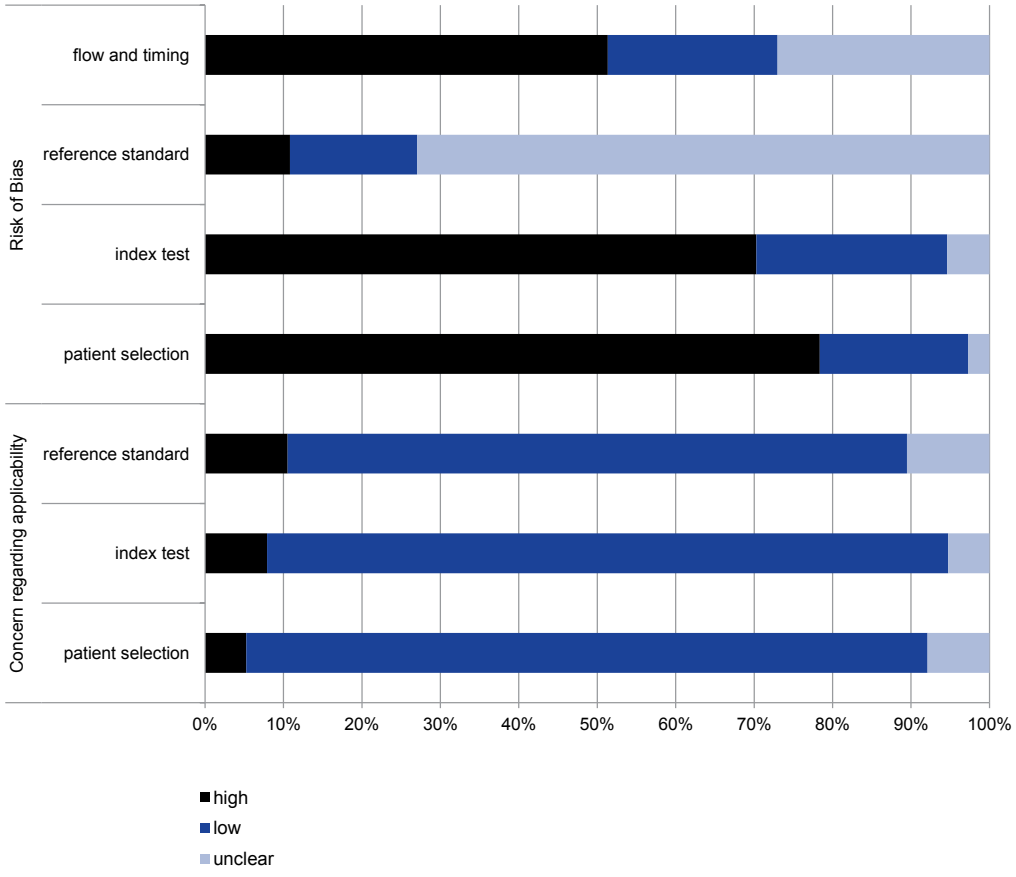
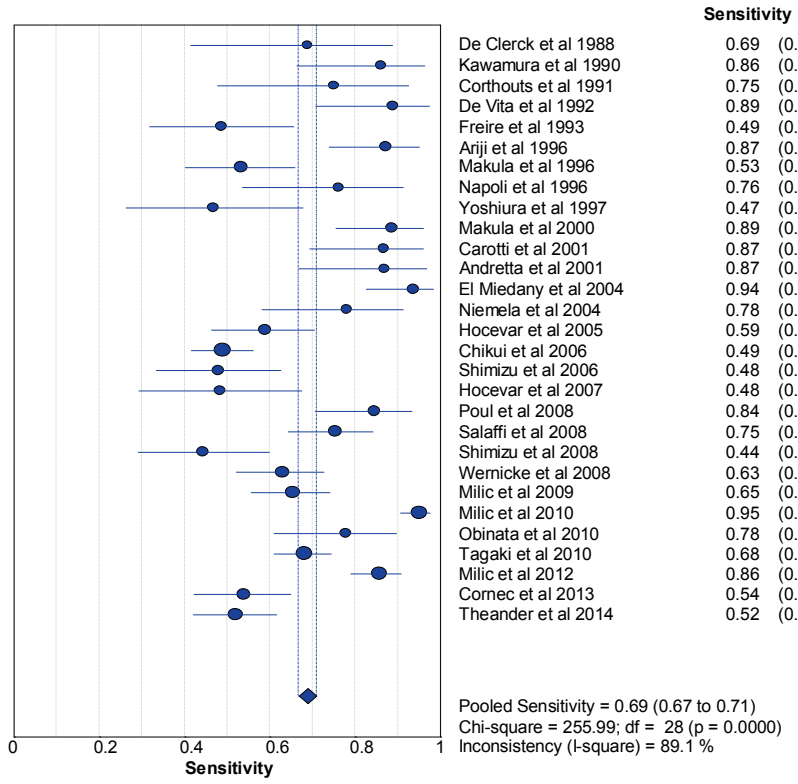


Figure 3: Forest plot showing study-specific and pooled sensitivity of ultrasonography in diagnosing Sjögren's syndrome. The point estimates of sensitivity for each study are shown as solid circles, and the size of each circle indicates the weight that it is given based on the sample size of each study. Error bars are 95% confidence intervals.



Quantitative synthesis

Publication bias

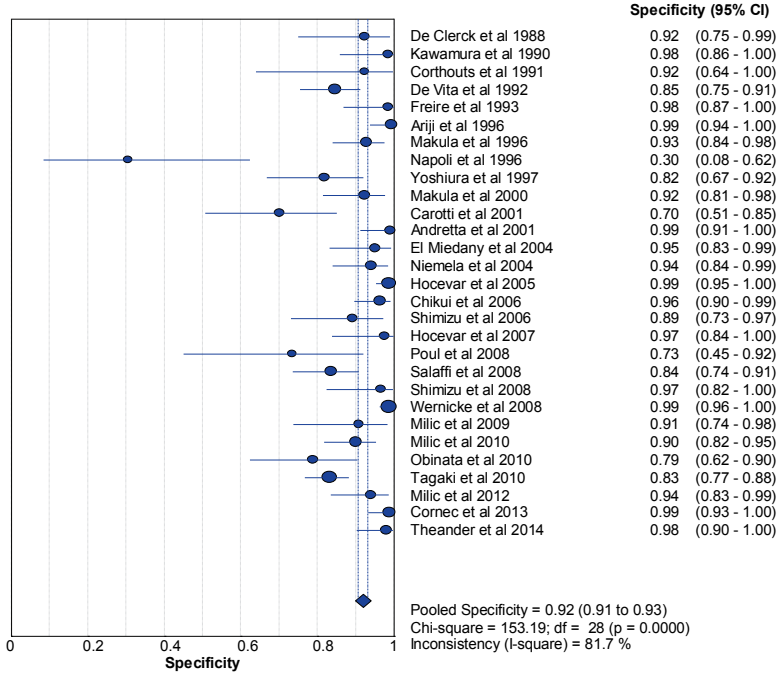
In the funnel plot, no studies were detected in the lower left part of the plot indicating publication bias (supplementary Figure 1).

Diagnostic accuracy of ultrasonography

In the meta-analysis, 29 studies using diagnostic criteria were included. Seven studies were excluded because: one study used only scintigraphy as reference standard [26], one study included sSS patients only [39], one study used only histology and sialography as reference standard [27], one study used only sialography as reference standard [20], and three studies had no control groups [16,44, 46]. Additionally, one study [40] was excluded because in that paper the study population was the same as in another paper [19]. We included only the paper with data on the number of true positives, true negatives, false negatives and false positives [19].

Regarding the diagnostic properties of ultrasonography in detecting SS, pooled sensitivity (Figure 3) was 0.69 (95% CI: 0.67 to 0.71) and pooled specificity (Figure 4) was 0.92 (95% CI: 0.91 to 0.93). We used a random effects model to determine the pooled diagnostic odds ratio (Figure 5): 33.89 (95% CI: 20.75 to 55.35). Considerable inconsistency was detected between studies when assessing sensitivity, specificity and diagnostic odds ratio (I^2= 89.1%, 81.7% and 72.4%).

Figure 4: Forest plot showing study-specific and pooled specificity of ultrasonography in diagnosing Sjögren's syndrome. The point estimates of specificity for each study are shown as solid circles, and the size of each circle indicates the weight that it is given based on the sample size of each study. Error bars are 95% confidence intervals.



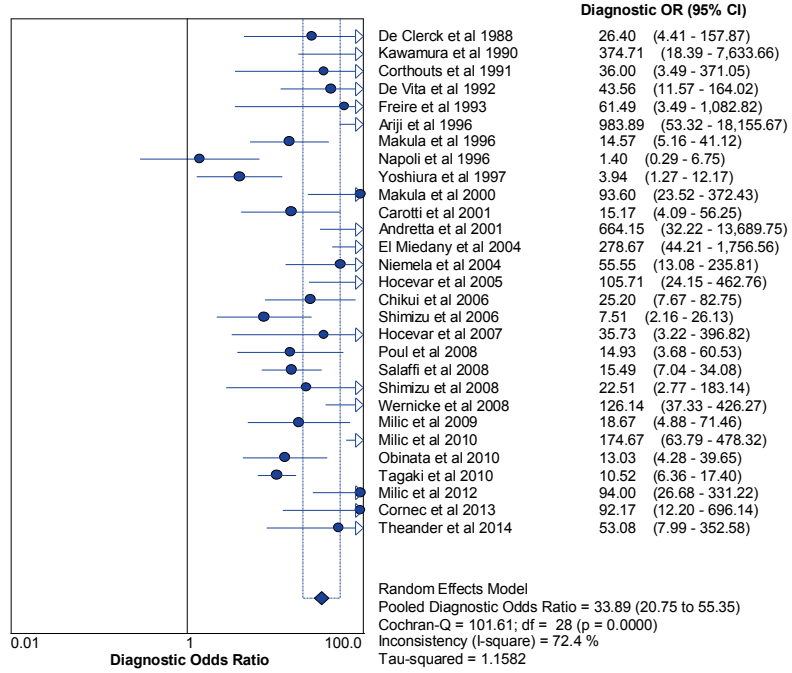
Meta-regression analysis

In the meta-regression analysis, it was found that studies with a high risk of bias in 'the conduct and interpretation of the index test' and studies using an ultrasonographic scoring system evaluating only homogeneity had a significantly higher natural logarithm of the diagnostic odds ratio (1.09 and 2.49, respectively; supplementary Table 4) in comparison to studies with a low risk of bias in the index test that used an ultrasonographic scoring system evaluating both ultrasonographic characteristics (echogenicity and homogeneity) for diagnosing SS.

Discussion

In the diagnosis of SS, involvement of salivary glands is currently assessed by sialography, scintigraphy, sialometry and histopathology. Recent discussion has focused on the accuracy of ultrasonography to evaluate the involvement of the major

Figure 5: Forest plot showing study-specific and pooled diagnostic odds ratio of ultrasonography in diagnosing Sjögren's syndrome. The point estimates of diagnostic odds ratio for each study are shown as solid circles, and the size of each circle indicates the weight that it is given based on the sample size of each study. Error bars are 95% confidence intervals. Calculations were performed using a random effects model. Dotted lines indicate whether the point estimate of each study falls within the confidence intervals of the pooled diagnostic odds ratio.



salivary glands and eventually to diagnose the disease [55]. Our study demonstrated that ultrasonography has a sensitivity, specificity and diagnostic odds ratio of 69%, 92% and 33.89, respectively, to diagnose SS in the major salivary glands. However, a high risk of bias in 'patient selection', 'index test' and 'flow and timing' was found in the included studies. Additionally, different ultrasonography scoring systems and study populations (study and control groups) were used in the studies, and publication bias was common. These methodological shortcomings probably inflated the diagnostic properties of ultrasonography in this meta-analysis. Additionally, studies with a poorer result (left lower part of the funnel plot) are missing (supplementary Figure 1). Several types of diagnostic criteria were used to identify SS patients in the included studies (Table 1), viz., AECG [47], ECSG [48], FC [49], RJDC [50], as well as less common ones like the Copenhagen Criteria [51], Takashima Criteria [43], and The Sjögren's Disease research Committee of the Japanese Ministry of Health and Welfare 1977 Diagnostic Criteria for definite Sjögren's Syndrome [52]. According to the meta-regression, the type of the cri-

teria used to diagnose SS did not seem to influence study outcome significantly. However, lack of power could explain this result.

The majority of included studies used a case control design. This means that the contrast in disease characteristics is large and that ultrasonography can 'easily' distinguish healthy controls from SS patients. The pooled specificity was high (92%), suggesting that ultrasonography may successfully identify patients who do not have SS. On the other hand, pooled sensitivity was considerably lower (69%), indicating that having a negative test is not adequate to exclude the presence of the disease.

Recently, a meta-analysis reported on the diagnostic properties of ultrasonography and sialography in SS [5]. In that study, 57 papers were initially identified and only 6 studies were finally included. The pooled sensitivity, specificity and diagnostic odds ratio was 77%, 81% and 17.48. In contrast to our study, the authors of that meta-analysis were less rigorous in their assessment of research methodology, since the number of studies with high risk of bias was limited in all QUADAS-2 domains. Concerns have been expressed regarding the outcome of their study [12], particularly the discrepancy between the data as included in that meta-analysis and the data provided by the source studies [6-11]. Additionally, the authors did not provide any data regarding inter-observer agreement.

Strengths and limitations

Strength of the current meta-analysis was the detailed literature search on eight databases, no language restriction, assessment of study eligibility by two reviewers, good inter-observer agreement, and application of the QUADAS-2 tool. The major limitation in the interpretation of the pooled outcomes (sensitivity, specificity and diagnostic odds ratio) is the low quality of the included studies and their clinical as well as their methodological heterogeneity. The likely sources of this heterogeneity are the variation in study populations, the ultrasonography scoring systems and the study designs. Another limitation in the interpretation of the pooled outcomes is publication bias, since small or large studies with positive results were overpresented in the funnel plot.

Implications and future research

We suggest that future studies should comply with the QUADAS-2 guidelines in order to ensure high diagnostic quality. Particular interest should be paid in the QUADAS-2 domains, where high risk of bias was observed: 1) a consecutive or random sample of patients should be used; a case control design and inappropriate exclusion of patients should be avoided, 2) ultrasonography results should be interpreted by observers blinded for each other as well as for the results of the reference test (diagnostic criteria, histology, sialography, scintigraphy, etc.); ideally, the applied threshold scoring should be pre-specified, 3) an appropriate and rather short interval should elapse between the application of ultrasonography and the reference test, the whole study population should receive the reference test (which should be always the same) and the whole study population should be included in the analysis.

The aforementioned features should be stated clearly by authors of the various papers to avoid potential misunderstanding and undervaluation of the study design. We acknowledge the need for a universally accepted ultrasonography scoring system to ensure uniform and standardized evaluation of the major salivary glands.

Conclusions

From the results of this meta-analysis we conclude that ultrasonography has the potential to evolve into a viable alternative in the evaluation of the major salivary glands in patients with SS, and therefore may be used as a non-invasive tool in the diagnosis of the disease. However, due to the low quality of the included studies, further research is required to elucidate the properties of ultrasonography in diagnosing SS.

Acknowledgements

The authors would like to thank Ms Sjoukje Van der Werf, biomedical information specialist in the University of Groningen, Groningen, the Netherlands, for her assistance in developing the search strategy. The authors would like to thank also Dr. Rada Moerman, Dr. Adriana Perez Fortis, Dr. Carina Boven and Dr. Wang Qi, research scientists in the University of Groningen, Groningen, the Netherlands, for translating the Russian, Spanish, Italian and Chinese articles, respectively.

Disclosure of conflict of interests

The authors state that they have no conflict of interests.

Funding

No funding has been received for the conduct of this study and/or preparation of this manuscript.

References

1. Vissink A, Bootsma H, Spijkervet FK, et al. Current and future challenges in primary Sjögren's syndrome. Curr Pharm Biotechnol 2012;13:2026-2045.

2. Kroese FGM and Bootsma H. Biomarkers: New biomarkers for Sjogren's Syndrome- time to treat patients. Nat Rev Rheumatol 2013;9:570-572.

3. Hocevar A, Ambrozic A, Rozman B, et al. Ultrasonographic changes of major salivary glands in primary Sjogren's syndrome. Diagnostic value of a novel scoring system. Rheumatology (Oxford) 2005;44:768-772.

4. Terkawi AS, Karakitsos D, Elbarbary M, et al. Ultrasound for the anesthesiologists: present and future. ScientificWorldJournal 2013: 683685.

5. Song GG and Lee YH. Diagnostic accuracies of sialography and salivary ultrasonography in Sjögren's syndrome patients: a meta-analysis. Clin Exp Rheumatol 2014;32:516-522.

6. Obinata K, Sato T, Ohmori K, et al. A comparison of diagnostic tools for Sjögren syndrome, with emphasis on sialography, histopathology, and ultrasonography. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2010;109:129-134.

7. Poul JH, Brown JE, Davies J. Retrospective study of the effectiveness of high-resolution ultrasound compared with sialography in the diagnosis of Sjogren's syndrome. Dentomaxillofac Radiol 2008:37:392-397.

8. Salaffi F, Carotti M, Iagnocco A, et al. Ultrasonography of salivary glands in primary Sjögren's syndrome: a comparison with contrast sialography and scintigraphy. Rheumatology (Oxford) 2008;47:1244-1249.

9. Takagi Y, Kimura Y, Nakamura H, et al. Salivary gland ultrasonography: can it be an alternative to sialography as an imaging modality for Sjögren's syndrome? Ann Rheum Dis 2010;69:20101321-1324.

10. Yonetsu K, Takagi Y, Sumi M, et al. Sonography as a replacement for sialography for the diagnosis of salivary glands affected by Sjögren's syndrome. Ann Rheum Dis 2002;61: 276-277.

11. Yoshiura K, Yuasa K, Tabata O, et al. Reliability of ultrasonography and sialography in the diagnosis of Sjögren's syndrome. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1997;83:400-407.

12. Delli K, Dijkstra PU, Spijkervet FKL, et al. Comment on 'Diagnostic accuracies of sialography and salivary ultrasonography in Sjögren's syndrome patients: a meta-analysis' by Song and Lee (2014). Clin Exp Rheumatol 2015;33:293.

13. Whiting PF, Rutjes AW, Westwood ME, et al. QUDAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529-536.

14. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009;151:264–269.

15. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.

16. Akin I, Esmer N, Gerceker M, et al. Sialographic and ultrasonographic analyses of major salivary glands. Acta Oto-Laryngol 1991;111:600-606.

17. Andretta M, Sfriso P, Botsios C, et al. Comparison of ultrasonography and sialography vs. magnetic resonance in the diagnosis of the primary Sjogren's syndrome. Acta Otorhinolaryngol Ital 2001;21:22-31.

18. Ariji Y, Ohki M, Eguchi K, et al. Texture analysis of sonographic features of the parotid gland in Sjogren's syndrome. AJR Am J Roentgenol 1996;166:935-941.

19. Carotti M, Salaffi F, Manganelli P, et al. Ultrasonography and colour Doppler sonography of salivary glands in primary Sjogren syndrome. Clin Rheumatol 2001;20:213-219.

20. Chikui T, Okamura K, Tokumori K, et al. Quantitative analyses of sonographic images of the parotid gland in patients with Sjogren's syndrome. Ultrasound Med Biol 2006;32:617-622.

21. Chikui T, Shimizu M, Kawazu T, et al. A quantitative analysis of sonographic images of the salivary gland: a comparison between sonographic and sialographic findings. Ultrasound Med Biol 2009;35:1257-1264.

22. Cornec D, Jousse-Joulin S, Pers JO, et al. Contribution of salivary gland ultrasonography to the diagnosis of Sjogren's syndrome: Toward new diagnostic criteria? Arthritis Rheum 2013;65:216-225.

23. Corthouts B, De Clerck LS, Francx L, et al. Ultrasonography of the salivary glands in the evalutation of Sjogren's syndrome. Comparison with sialography. J Belge Radiol 1991;74:189-192.

24. De Clerck LS, Corthouts R, Francx L, et al. Ultrasonography and computer tomography of the salivary glands in the evaluation of Sjogren's syndrome. Comparison with parotid sialography. J Rheumatol 1988;15:1777-1781.

25. De Vita S, Lorenzon G, Rossi G, et al. Salivary gland echography in primary and secondary Sjogren's syndrome. Clin Exp Rheumatol 1992;10:351-356.

26. Decuzzi M, Tatulli F, Giampaolo M, et al. Sialocintigraphy versus ultrasonography of the salivary glands in patients first diagnosed with Sjogren's syndrome. Hell J Nucl Med 2006;9:103-105.

27. Diederich S, Wernecke K, Peters PE. Sialographic and sonographic diagnosis of salivary gland diseases. Radiologe 1987;27:255-261.

28. El Miedany YM, Ahmed I, Mourad HG, et al. Quantitative ultrasonography and magnetic resonance imaging of the parotid gland: can they replace the histopathologic studies in patients with Sjogren's syndrome? Joint Bone Spine 2004;71:29-38.

29. Freire Gonzalez M, Grana Gil J, Martinez Sapina MJ, et al. The performance of echography of the salivary glands in the Sjogren's syndrome and its concordance with other diagnostic tests. Rev Esp Reumatol 1993;20:75-79.

30. Hocevar A, Rainer S, Rozman B, et al. Ultrasonographic changes of major salivary glands in primary Sjogren's syndrome - Evaluation of a novel scoring system. Eur J Radiol 2007;63: 79-383.

31. Kawamura H, Taniguchi N, Itoh K, et al. Salivary gland echography in patients with Sjogren's syndrome. Arthritis Rheum 1990;33:505-510.

32. Makula E, Pokorny G, Kiss M, et al. The place of magnetic resonance and ultrasonographic examinations of the parotid gland in the diagnosis and follow-up of primary Sjögren's syndrome. Rheumatology (Oxford) 2000;39:97-104.

33. Makula E, Pokorny G, Rajtar M, et al. Parotid gland ultrasonography as a diagnostic tool in primary Sjogren's syndrome. Br J Rheumatol 1996;35:972-977.

34. Milic V, Petrovic R, Boricic I, et al. Ultrasonography of major salivary glands could be an alternative tool to sialoscintigraphy in the American-European classification criteria for primary Sjogren's syndrome. Rheumatology (Oxford) 2012;51:1081-1085.

35. Milic VD, Petrovic RR, Boricic IV, et al. Diagnostic Value of Salivary Gland Ultrasonographic Scoring System in Primary Sjogren's Syndrome: A Comparison with Scintigraphy and Biopsy. J Rheumatol 2009;36:1495-1500.

36. Milic VD, Petrovic RR, Boricic IV, et al. Major salivary gland sonography in Sjogren's syndrome: Diagnostic value of a novel ultrasonography score (012) for parenchymal inhomo-

geneity. Scand J Rheumatol 2010;39:160-166.

37. Napoli V, Tozzini A, Neri E, et al. The imaging diagnosis of Sjogren's syndrome: echography, sialography and scintigraphy compared in the study of the salivary glands. Minerva Stomatol 1996;45:141-148.

38. Niemela RK, Takalo R, Paakko E, et al. Ultrasonography of salivary glands in primary Sjogren's syndrome. A comparison with magnetic resonance imaging and magnetic resonance sialography of parotid glands. Rheumatology (Oxford) 2004;43:875-879.

39. Nitsche A, Pérez Dávila A, Seleme G, et al. Ecografía de glándulas submandibulares y parótida en Sindrome de Sjögren. Nuestra Experiencia/Echography saluvary gland submandibular and parotid in Sjögren Syndrome. Rev argent reumatol 1996;7:5-9.

40. Salaffi F, Argalia G, Carotti M, et al. Salivary gland ultrasonography in the evaluation of primary Sjogren's syndrome. Comparison with minor salivary gland biopsy. J Rheumatol 2000;27:1229-1236.

41. Shimizu M, Okamura K, Yoshiura K, et al. Sonographic diagnosis of Sjogren syndrome: evaluation of parotid gland vascularity as a diagnostic tool. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2008;106:587-594.

42. Shimizu M, Okamura K, Yoshiura K, et al. Sonographic diagnostic criteria for screening Sjogren's syndrome. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2006;102:85-93.

43. Takashima S, Morimoto S, Tomiyama N, et al. Sjogren syndrome: comparison of sialography and ultrasonography. JCU 1992;20:99-109.

44. Theander E and Mandl T. Primary Sjogren's syndrome: The diagnostic and prognostic value of salivary gland ultrasonography using a simplified scoring system. Arthritis Care Res (Hoboken) 2013;66:1102-1107.

45. Varshavskii AI, Panchenko KI, Guberskaia TA, et al. Comparative sonographic, x-ray and morphological studies of the salivary glands in Sjogren's syndrome. Ter Arkh 1999;71:61-67.

46. Wernicke D, Hess H, Gromnica-Ihle E, et al. Ultrasonography of salivary glands - a highly specific imaging procedure for diagnosis of Sjogren's syndrome. J Rheumatol 2008;35:285-293.

47. Vitali C, Bombardieri S, Jonsson R, et al. Classification criteria for Sjögren's syndrome: a revised version of the European criteria proposed by the American-European Consensus Group. Ann Rheum Dis 2002;61:554-558.

48. Vitali C, Bombardieri S, Moutsopoulos HM, et al. Preliminary criteria for the classification of Sjögren's syndrome. Results of a prospective concerted action supported by the European Community. Arthritis Rheum 1993;36:340-347.

49. Fox RI, Robinson CA, Curd JG, et al. Sjögren's syndrome. Proposed criteria for classification. Arthritis Rheum 1986;29:577-585.

50. Miyawaki S. Revised Japan criteria for Sjögren syndrome. Ryumachi 2000;40:48-53.

51. Manthorpe R, Oxholm P, Prause JU, et al. The Copenhagen criteria for Sjögren's syndrome. Scand J Rheumatol Suppl 1986;61:19-21.

52. Ohfuji T. Annual report of the Ministry of health and Welfare: The Sjögren's Disease research Committee. Japan 1977 (in Japanese).

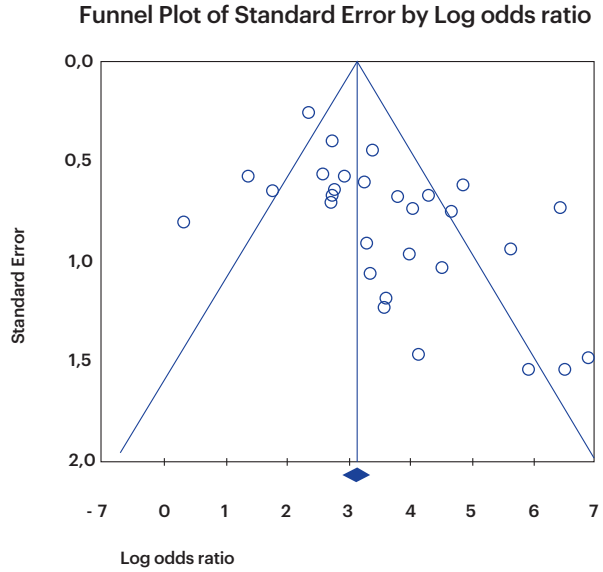
53. Wittich GR, Scheible WF, Hajek PC. Ultrasonography of the salivary glands. Radiol Clin North Am. 1985;23:29-37.

54. Da-Xi S, Hai-Xiong S, Qiang Y. The diagnostic value of ultrasonography and sialography in salivary gland masses. Dentomaxillofac Radiol 1987;16:37-45.

55. Bootsma H, Spijkervet FK, Kroese FG, et al. Toward new classification criteria for Sjögren's syndrome? Arthritis Rheum 2013;65: 21-23.

Supplementary files

Supplementary Figure: Funnel plot of standard error by log odds ratio. No studies were detected in the lower left part, indicating that publication bias is very likely.



Supplementary Table 1: Electronic databases and search details (according to the syntax rules of each database).

Database	Search Details
PubMed (http://www.ncbi.nlm.nih.gov/pubmed/)	("Sjogren's Syndrome"[Mesh] OR sjogren*[tw] OR sicca syndrome[tw])AND("Ultrasonography"[Mesh] OR "ultrasonography" [Subheading] OR ultraso*[tw] OR echograph*[tw] OR echotomograph*[tw] OR sonograph*[tw])
EMBASE (http://www.embase.com/home)	(Sjogren* OR 'sicca syndrome') AND (ultraso* OR echograph* OR echotomograph* OR sonograph*)
LILACS (bases.bvs.br)	(tw:((Sjogren* OR "sicca syndrome"))) AND (tw:((ultraso* OR echograph* OR echotomograph* OR sonograph*)))
Web of Science (scientific.thomson.com/products/wos/)	(Sjogren* OR "sicca syndrome") AND (ultraso* OR echograph* OR echotomograph* OR sonograph*)
OpenGrey (http://www.opengrey.eu/)	sjogren* AND ultrasound
African Journals online (www.ajol.info)	sjogren*
Cochrane Central Register of Controlled Trials (www.thecochranelibrary.com)	(Sjogren* OR "sicca syndrome") AND (ultraso* OR echograph* OR echotomograph* OR sonograph*)
National Institutes of Health Clinical Trials Registry (clinicaltrials.gov)	Sjogren AND ultrasound [Recognized Terms and Synonyms: ultrasound: 3947 studies, (echography, echotomography, gen.us, sonogram, sonography, ultrasonic, ultrasonic imaging, ultrasonic shockwave, ultrasonogram, ultrasonography), sjogren: 85 studies (gougerot-nulock-houwer syndrome, sjoegren syndrome, syndrome sjogren's)]

Supplementary Table 2: Outcomes of the risk of bias assessment and concerns regarding applicability of studies included in the quality assessment.

		Risk of Bias				Concern regarding applicability		
Publication Year	Author	f/t	rs	it	ps	rs	It	ps
1987	Diederich et al	high	unclear	high	high	unclear	low	low
1988	De Clerck et al	high	unclear	high	high	high	high	high
1990	Kawamura et al	high	unclear	high	high	low	high	low
1991	Akin et al	high	unclear	high	high	low	low	low
1991	Corthouts et al	high	unclear	high	low	low	low	low
1992	De Vita et al	unclear	unclear	high	high	low	low	low
1992	Takashima et al	high	unclear	high	high	unclear	low	unclear
1993	Freire et al	low	high	low	high	high	low	low
1996	Ariji et al	high	low	unclear	high	low	low	low
1996	Makula et al	high	unclear	high	high	low	low	low
1996	Napoli et al	unclear	unclear	high	high	low	low	low
1996	Nitsche et al	low	high	low	unclear	unclear	low	unclear
1997	Yoshiura et al	high	unclear	high	high	low	low	low
1999	Varshavsky	unclear	unclear	high	high	unclear	unclear	unclear
2000	Makula et al	high	high	high	high	low	low	low
2000	Salaffi et al	unclear	unclear	unclear	high	low	low	low
2001	Andretta et al	unclear	unclear	high	high	low	low	low
2001	Carotti et al	low	low	low	high	low	low	low
2004	El Miedany et al	high	low	high	high	low	low	low
2004	Niemela et al	high	unclear	high	high	low	low	low
2005	Hocevar et al	high	unclear	high	low	high	low	low
2006	Chikui et al	low	low	high	low	low	low	low
2006	Decuzzi et al	unclear	unclear	high	high	high	low	high
2006	Shimizu et al	unclear	unclear	high	high	low	low	low
2007	Hocevar et al	high	low	low	high	low	low	low
2008	Poul et al	high	unclear	high	high	low	low	low
2008	Salaffi et al	low	low	low	high	low	low	low
2008	Shimizu et al	high	unclear	low	high	low	low	low
2008	Wernicke et al	high	unclear	high	high	low	low	low
2009	Chikui et al	low	high	low	low	low	low	low
2010	Milic et al	high	unclear	high	high	low	low	low
2010	Obinata et al	high	unclear	low	high	low	low	low
2010	Tagaki et al	unclear	unclear	low	high	low	low	low
2012	Milic et al	low	unclear	high	low	low	low	low
2013	Cornec et al	unclear	unclear	high	low	low	high	low
2013	Milic et al	low	unclear	high	low	low	low	low
2014	Theander et al	unclear	unclear	high	high	low	low	low

f/t: flow of patients and timing of performance of index test and reference standard, it: conduct and interpretation of index test, ps: patient selection, rs: conduct and interpretation of reference standard.

Supplementary Table 3: Summary of ultrasonographic scoring systems for Sjögren's syndrome applied in the different studies.

Author	Year	US scoring system
Diederich et al	1987	1) Homogenicity 2) Echogenicity (compared to thyroid gland) 3) Size (normal: parotid gland 40x15mm and SM 30x15mm)
De Clerck et al	1988	2-grade system: decreased echogenicity and normal echogenicity (compared to the masseter muscle and mylohyoid muscle)
Kawamura et al	1990	2-grade system: Homogenous parenchyma and heterogeneous parenchyma
Akin et al	1991	Evaluation based on reports of: Wittich et al 1985 [53] and Da-Xi et al 1987 [54]
Corthouts et al	1991	2-grade system: decreased reflectivity and normal reflectivity (compared to the masseter muscle and mylohyoid muscle)
De Vita et al	1992	1) Parenchymal inhomogenicity (mild/evident/gross) 2) Parenchymal echogenicity (increased/decreased) 3) Glandular volume (arbitrary evaluated; increased/decreased) 4) Posterior glandular border (definite/ill defined/not visible) 5) Lymph nodes (evidence of peri-glandular/evidence of intra-glandular)
Takashima et al	1992	US was positive for SS if multiple hypoechoic areas were found and intermediate on the presence of non-homogenous parotid glands
Freire et al	1993	Heterogeneity of the parenchyma and decrease of the size of the glands (parotid and submandibular) was considered as positive for SS.
Ariji et al	1996	A) 5 grade-system: 0= regular contour, no hypoechoic spots/areas, no echogenic bands 1= regular contour, small hypoechoic spots/areas, no echogenic bands 2= regular contour, round multiple hypoechoic spots/areas, no echogenic bands 3= irregular contour, round multiple hypoechoic spots/areas, presence of echogenic bands 4= irregular contour, irregular multiple hypoechoic spots/areas, presence of echogenic bands B) Quantitative characteristics: S value and SD value
Makula	1996	1) Parenchymal inhomogenicity (mild/evident/gross) 2) Parenchymal echogenicity compared to masseter (increased/decreased) 3) Glandular size (Parotid gland was normal if width= 27±7mm)
Napoli et al	1996	1) Ultrasonography of parotid gland: negative=0, positive=1 2) Volume: normal=0, increased=1 3) Border: normal=0, abnormal=1 4) Structure: homogeneous=0, non-homogeneous=1 5) Hypoechogenic areas: absence=0, <1cm=1, >1cm=2 6) Ductal ecstasia: absent=0, 2mm=1, 2-4mm=2, 4mm=3 7) Lymph nodes: absence=1, presence=1
Nitsche et al	1996	1) parenchyma echogenicity (homogenous or non-homogenous) 2) size of gland (longitudinal and transverse diameter) 3) presence/absence of lymph nodes
Yoshiura et al	1997	Five-point-rating scale: 1: definitely normal/ 2: probably normal/ 3: not sure/ 4: probably abnormal/ 5: definitely abnormal
Makula et al	2000	1) parenchymal inhomogeneity (PIH), three grades of PIH were distinguished. In mild PIH (grade 1), a diffuse microareolar structure can be seen, the borders of the hypoechogenic areolae are blurred, and the areolae are <2 mm in diameter. In evident (moderately severe) PIH (grade 2), the hypoechogenic areas are larger (2–6 mm in diameter), with a sharper border. In gross (severe) PIH (grade 3), large (>6 mm in diameter) circumscribed hypoechogenic areas are also present 2) The parenchymal echogenicity was determined in comparison with that of the thyroid gland 3) The size of the parotid was considered to be normal if its width was 27±7 mm Two groups of SS patients were differentiated on the basis of the US findings: patients with a homogeneous parotid gland parenchyma and mild PIH (grade 1), and patients with more advanced abnormalities (grade 2 or 3) which are of true diagnostic value
Salaffi et al	2000	-Grade 0: normal gland -Grade 1: regular contour, small hypoechoi spots/areas without echogenic bands, regular or increased glandular volume (mean values 20±3mm for the parotids and 13±2mm for the submandibular glands), and ill defined posterior glandular border -Grade2: regular contour, evident multiple scattered hypoechogenic areas usually of variable size, (<2mm) and not uniformly distributed, without echogenic bands, regular or increased glandular volume and ill defined posterior glandular border -Grade 3: irregular contour, multiple large circumscribed or confluent hypoechogenic areas (2-6mm) and/or multiple cysts, with echogenic bands, regular or decreased glandular volume and posterior glandular border not visible -Grade 4: irregular contour, multiple large circumscribed or confluent hypoechogenic areas (>6mm) and/or multiple cysts or multiple calcifications, with echogenic bands, resulting in severe damage to the glandular architecture, decreased glandular volume, and posterior glandular border not visible Score ranges from 0-16 (comprising of the sums of the single scores 0-4 for each parotid and submandibular gland Optimal grade cutoff=8
Andretta et al	2001	Grade 0: no alterations Grade 1: gland increased in size or with normal echogenicity of the most reflective background Grade 2: gland increased in size or normal with sinuous profiles and inhomogeneous echogenicity with aspects of microareolara Grade 3: gland increased in size with sinuous profiles, inhomogeneous echogenicuty and coarse areola Grade 4: as above with ecstatic ducts and microcalculi in gland or small, sinuous profiles and grossly inhomogeneous echogenicity (type atrophic gland)

Supplementary Table 3 (continued)

Author	Year	US scoring system
El Miedany et al	2004	Grade 0: Normal homogenous parenchyma Grade 1: Mild PIH seen as diffuse hypoechoic areolae less than 2 mm with blurred borders Grade 2: Moderate PIH seen as large hypoechoic areas, 2–6 mm diameter, with sharp borders Grade 3: Severe PIH with large, more than 6 mm circumscribed hypoechoic areas
Niemela et al	2004	-stage 0=normal parenchymal structure; -stage 1=mild parenchymal inhomogeneity (PIH) (hypoechoic areas <2 mm) -stage 2=evident PIH (hypoechoic areas of 2–6 mm) -stage 3=gross PIH (hypoechoic areas >6 mm); and -stage 4=adipose degeneration of the gland (adipose tissue echogenicity and parenchymal atrophy)
Hocevar et al	2005	1) Parenchymal echogeneity was evaluated in comparison with the thyroid gland or when there was coincident thyroid gland disease by surrounding anatomical structures (muscular structures, subcutaneous fat). If the echogeneity was comparable to the thyroid, the grade was 0; if it was decreased, we graded it 1 2) Homogeneity was graded from 0 to 3. Grading 0 was for a homogeneous gland, 1 for mild inhomogeneity, 2 for evident inhomogeneity, and 3 for a grossly inhomogeneous gland 3) The presence of hypoechogenic areals was graded from 0 to 3 (grade 0, absent; grade 1, a few, scattered; grade 2, several; grade 3, numerous hypoechogenic areas) 4) Hyperechogenic reflections were graded from 0 to 3 in the parotid glands (grade 0, absent; grade 1, a few, scattered; grade 2, several; grade 3, numerous hyperechogenic reflections) and from 0 (absent) to 1 (present) in the submandibular glands. 5) Clearness of salivary gland borders was graded from 0 to 3 (grade 0, clear, regular defined borders; grade 1, partly defined borders; grade 2, ill-defined borders; grade 3, borders not visible) Final score: summation of the grades for the five parameters described above for all four glands.US score ranged from 0 to 48 Optimal grade cutoff=17
Chikui et al	2006	A) 3-grade system +: definitely present, ±: probably present,-: definitely absent B) Quantitative characteristics: Standard deviation of regions of interest (SD), Angular second moment (ASM), Correlation (COR), Inverse difference moment (IDM), Entropy (ENT), Hurst co-efficient (H)
Decuzzi et al	2006	0= normal I= early destructive changes (regular border, normal or enlarged dimensions and puntate hypoechoic lesions) II= late destructive changes (irregular border, reduced dimensions, multicystic or reticular pattern, calcifications and hyperechoic stripes)
Shimizu et al	2006	- multiple hypoechoic areas - multiple hyperechoic lines and/or spots - multiple hypoechoic areas surrounded with hyperechoic lines and/or spots On the sonographic images of the submandibular glands obscuration of the gland configuration was also evaluated additionally. Findings in the parotid gland are designated "P" and in the submandibular gland "S". A total of 7 findings were evaluated: P1, P2, P3, S1, S2, S3, and S4 combining both types of gland. Evaluations were recorded using 3 grades: positive, probable, and negative.
Poul et al	2008	US was considered positive for SS if the following features were detected: bilateral decreased parotid gland reflectivity, and heterogeneous or nodular parenchyma with a honeycomb appearance.
Salaffi et al	2008	Salaffi et al 2000, BUT: Optimal grade cutoff=6

Supplementary Table 4: Results of the meta-regression (method of moments) to analyze effects of study methodology on the natural logarithm of the diagnostic odds ratio using a random effects model.

Study characteristics	Regression Coefficient	95% CI	P
Intercept (reference study)	2.46	1.65 - 3.27	< 0.001
High risk of bias in index test	1.09	0.17 - 2.02	0.021
US characteristic: echogenicity assessed only	0.10	-0.94 - 1.15	0.845
US characteristic: homogeneity assessed only	2.49	0.82 - 4.15	0.003

95% CI: 95% confidence interval, reference study is a study with: a low risk of bias in the index test using both ultrasonographic characteristics (echogenicity and homogeneity) for diagnosing SS. High risk of bias in index test: the conduct and interpretation of the index test is highly likely to be biased. (Tau^2 = 0.6661, I^2 = 58.98%, Q = 60.95, df = 25, p = 0.0001). Other potential predictors were not significantly related to diagnostic properties of ultrasound in SS found in the studies.